

本文引文格式:李菲,王宇航,徐陈晨,等. EDNRB 在前列腺癌中的表达及临床预后意义[J]. 右江民族医学院学报, 2021, 43(5): 611-616.

【论著与临床报道】

EDNRB 在前列腺癌中的表达及临床预后意义

李菲¹, 王宇航¹, 徐陈晨¹, 张永静¹, 李伟玉¹, 陆进^{2,3}

(1. 蚌埠医学院临床医学院, 安徽 蚌埠 233030;

2. 蚌埠医学院人体解剖学教研室, 安徽 蚌埠 233030;

3. 蚌埠医学院数字医学与智慧健康安徽省重点实验室, 安徽 蚌埠 233030)

摘要:目的 利用生物信息学方法, 筛选前列腺癌预后的关键基因。方法 首先, 从 GEO 数据库下载 4 个前列腺癌表达谱芯片数据集(GSE3325/GSE46602/GSE55945/GSE69223); 其次, 分别利用 GEO2R、DAVID 以及 STRING 数据库分析前列腺癌的差异表达基因(DEGs), 并对差异基因进行基因本体论(GO)和京都基因与基因组百科全书(KEGG)通路富集以及蛋白质-蛋白质相互作用关系(PPI); 再次, 通过 Cytoscape 软件的 Cytohubba 插件 EPC 算法筛选关键基因; 最后, 利用基因表达谱交互分析(GEPIA)数据库对关键基因进行预后与表达分析, 利用人类蛋白图谱(HPA)数据库下载 TCGA 数据对结果进行验证。结果 筛选出 103 个 DEGs, 其功能主要富集在血管生成、脂质代谢以及细胞分化等方面, 而通路主要富集在 Hippo 信号上。筛选的 10 个关键基因中仅 EDNRB 对前列腺癌患者的预后具有显著影响, 且其低表达不利于患者预后。结论 通过生物学分析筛选前列腺癌关键基因, 为其治疗提供新的分子靶点。

关键词: 前列腺肿瘤; 计算生物学; 基因; 总体生存; 预后

中图分类号: R737.25

文献标识码: A

文章编号: 1001-5817(2021)05-0611-06

doi:10.3969/j.issn.1001-5817.2021.05.008

Expression and prognostic significance of EDNRB in prostate cancer

Li Fei¹, Wang Yuhang¹, Xu Chenchen¹, Zhang Yongjing¹, Li Weiyu¹, Lu Jin^{2,3}

(1. School of Clinical Medicine, Bengbu Medical College, Bengbu 233030, Anhui, China;

2. Department of Human Anatomy, Bengbu Medical College, Bengbu 233030, Anhui,

China; 3. Anhui Key Laboratory of Computational Medicine and Intelligent Health,

Bengbu Medical College, Bengbu 233030, Anhui, China)

Abstract: **Objective** To screen the key genes for prognosis of prostate cancer by bioinformatics. **Methods** Firstly, four microarray datasets (GSE3325/GSE46602/GSE55945/GSE69223) of prostate cancer expression profile were downloaded from the GEO database. Secondly, the GEO2R, DAVID and STRING databases were used to analyze differentially expressed genes (DEGs) in prostate cancer. Then the pathway enrichment of gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) and the protein-protein interaction (PPI) for DEGs were analyzed. Thirdly, Cytohubba plug-in EPC algorithm of software Cytoscape was used to screen key genes. Finally, the gene expression profile interaction analysis (GEPIA) database was used to analyze the prognosis and expression of key genes, and the results were verified using data downloaded from the Human Protein Atlas (HPA) database. **Results** A total of 103 DEGs were screened out, and their functions were mainly enriched in angiogenesis, lipid metabolism and cell differentiation, while their pathways were mainly enriched in Hippo signals. Among the 10 key genes screened, only EDNRB had a significant effect on the prognosis of patients with prostate cancer, and its low expression was not conducive to the prognosis of patients. **Conclusion** Screening key genes of prostate cancer by bioinformatics can provide new molecular targets for its treatment.

Key words: prostate tumor; computational biology; gene; overall survival; prognosis

基金项目:安徽省大学生创新训练项目(S202010367071);教育部产学合作协同育人项目(202101160001)

第一作者简介:李菲(2000-),女,在读本科生, E-mail:1191091835@qq.com

通讯作者简介:陆进(1987-),男,在读硕士研究生,实验师,研究方向:肿瘤生物信息学, E-mail:0100197@bbmc.edu.cn

前列腺癌(prostate cancer, PCa)是最常见影响男性健康的一种恶性生殖系统肿瘤,是全球范围内男性癌症中的第二大死因^[1]。截至 2017 年,美国前列腺癌的确证病例超 19 万余,死亡病例超 3.3 万余人^[2]。在中国前列腺癌的发病率较西方国家低,但是由于生活方式改变等原因,近年来开始呈现明显的上升趋势^[3]。前列腺癌病因复杂,目前未被完全阐明,但遗传因素和基因改变目前被认为是其最关键的风险因素^[4]。近些年来,利用生物信息学分析癌症发生发展过程中出现的遗传或表观遗传学改变,并寻找与癌症诊断和预后相关的生物标志物是一种新的方法^[5]。本研究通过生物信息挖掘技术,利用多个数据库进行数据分析,来挖掘前列腺癌潜在的生物标记物,为其诊断治疗提供新的策略。

1 资料与方法

1.1 数据来源 GEO 数据库(<http://www.ncbi.nlm.nih.gov/geo>)是一个包含了高通量基因表达数据集、基因芯片和微阵列的数据库。从 GEO 数据库中筛选并下载 4 组前列腺癌数据集 GSE3325^[6]、GSE46602^[7]、GSE55945^[8]、GSE69223^[9],见表 1。

表 1 数据集具体信息

芯片名称	正常样本量	肿瘤样本量
GSE3325	6	13
GSE46602	10	36
GSE55945	8	13
GSE69223	15	15

1.2 差异表达基因的筛选 利用 GEO2R 对 4 组前列腺癌(PCa)数据进行分析,并通过绘制韦恩图筛选 PCa 的差异表达基因(DEGs),筛选标准为 $P < 0.01$ 且 $|\log_2FC| \geq 1$,去除没有对应基因符号的探针组或同一基因对应多个探针等情况。

1.3 GO 和 KEGG 富集分析 DAVID 数据库(<https://david.ncifcrf.gov/>)中包括生物数据和分析工具,可以提供基因和蛋白质信息注释功能。基因本体论(Gene Ontology,GO)功能注释是一种分析注释基因生物学过程的生物信息学工具。京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes,KEGG)通路富集分析是一个从分子水平来了解生物系统高级功能的数据库。利用 DAVID 对 DEGs 进行 GO 和 KEGG 富集分析,筛选条件: $P < 0.05$ 且 $Count \geq 5$ 。

1.4 蛋白互作网络及关键基因筛选 STRING 数据库(<http://string-db.org>)可以用来预测蛋白质-蛋白质相互作用关系(protein-protein interactions, PPI)。

利用 STRING 数据库对 DEGs 进行 PPI 网络分析,并将数据导入 Cytoscape 软件中,利用 Cytoscape 中的 Cytohubba 插件的 EPC 算法计算得分前 10 的基因作为关键基因。

1.5 关键基因的表达与生存预后分析 基因表达谱动态分析(Gene Expression Profiling Interactive Analysis,GEPIA)是一种成熟的可以用于分析 RNA 测序表达数据的网页工具。通过 GEPIA 分析关键基因在前列腺癌组织以及正常前列腺组织中的预后和表达情况。人类蛋白图谱数据库(Human Protein Atlas, HPA)是一种从 RNA 和蛋白水平研究人类不同组织和器官中的蛋白表达情况的网页工具。通过 HPA 下载 TCGA 数据作图验证 GEPIA 的分析结果。

2 结果

2.1 筛选差异表达基因 对 4 组数据集以 $P < 0.01$ 且 $|\log_2FC| \geq 1$ 为条件处理后,进行火山图和韦恩图绘制,获得 4 组公共数据共有的 103 个差异表达基因,其中上调基因 34 个,下调基因 69 个(见图 1、图 2)。

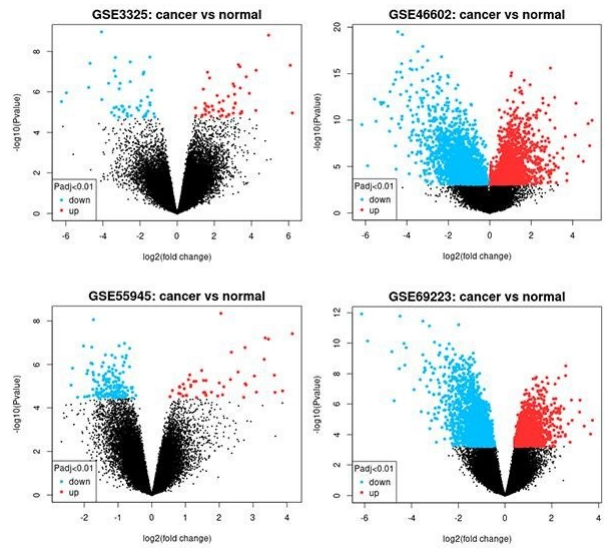


图 1 火山图

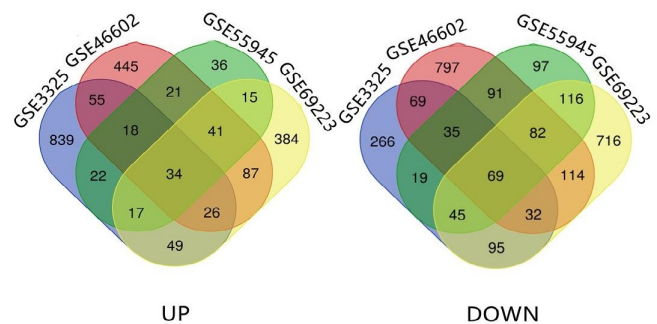


图 2 韦恩图

2.2 GO 功能注释和 KEGG 通路富集分析 使用 DAVID 数据库对 103 个 DEGs 进行 GO 和 KEGG 富

集分析(见表 2)。GO 分析显示,DEGs 的细胞学过程(biological processes,BP)主要包括 RNA 聚合酶 II 启动子转录负调控、血管生成、脂质代谢、细胞分化过程等;分子功能(molecular function,MF)主要包括蛋白

质异二聚体活性等;细胞成分(cell component,CC)主要包括细胞外基质、细胞质的核外区域、质膜等相关。KEGG 通路分析显示差异基因主要与局部黏附性、Hippo 信号通路显著相关。

表 2 差异基因的 GO 分析和 KEGG 分析

类别	项目	描述	数量	P	
BP	GO:000122	negative regulation of transcription from RNA polymerase II promoter	13	4.04E-04	
	GO:0001525	angiogenesis	6	0.00664572	
	GO:0006629	lipid metabolic process	5	0.009711128	
	GO:0007010	cytoskeleton organization	5	0.01058009	
	GO:0030154	cell differentiation	8	0.011330791	
	GO:0001666	response to hypoxia	5	0.013224713	
	GO:0007155	cell adhesion	7	0.035062826	
MF	GO:0046982	protein heterodimerization activity	7	0.032905199	
CC	GO:0005578	proteinaceous extracellular matrix	9	6.82E-05	
	GO:0005901	caveola	5	3.39E-04	
	GO:0031012	extracellular matrix	8	7.99E-04	
	GO:0005925	focal adhesion	9	8.70E-04	
	GO:0005788	endoplasmic reticulum lumen	5	0.016851244	
	GO:0030018	Z disc	5	0.003130742	
	GO:0005576	extracellular region	17	0.00715193	
	GO:0005886	plasma membrane	33	0.00621592	
	GO:0031965	nuclear membrane	6	0.006366659	
	GO:0048471	perinuclear region of cytoplasm	9	0.01401035	
	GO:0045121	membrane raft	5	0.021215334	
	GO:0031410	cytoplasmic vesicle	5	0.032298412	
	KEGG	hsa04510	Focal adhesion	5	0.038805777
		hsa04390	Hippo signaling pathway	5	0.014174875

2.3 构建 PPI 网络筛选关键基因 利用 STRING 数据库对 103 个 DEGs 进行分析(见图 3),通过 Cyto-scape 软件的 Cytohubba 插件的 EPC 算法筛选,得到

CAV1、SNAI2、TGFB3、CAV2、PGR、MYOF、EHD2、FGFR2、EDNRB、PRKG110 个关键基因(见图 4)。

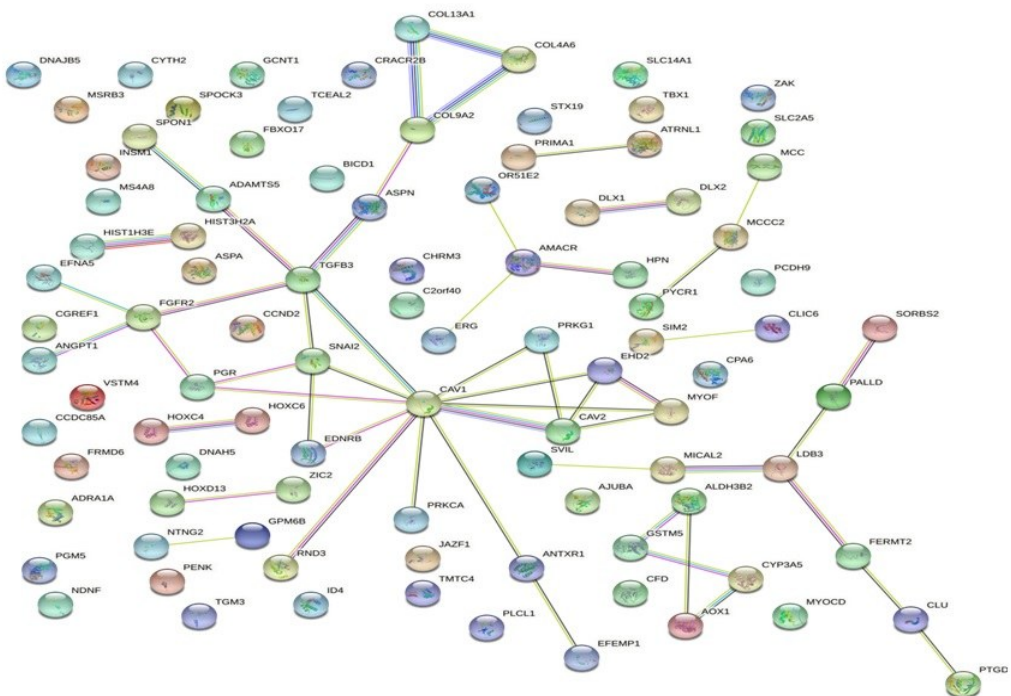


图 3 PPI 网络

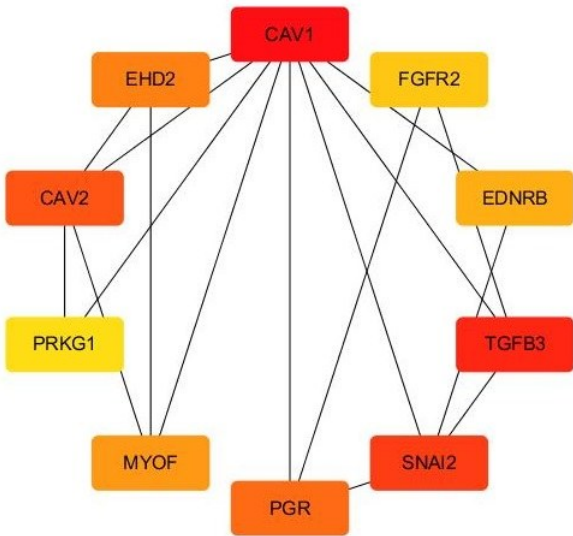
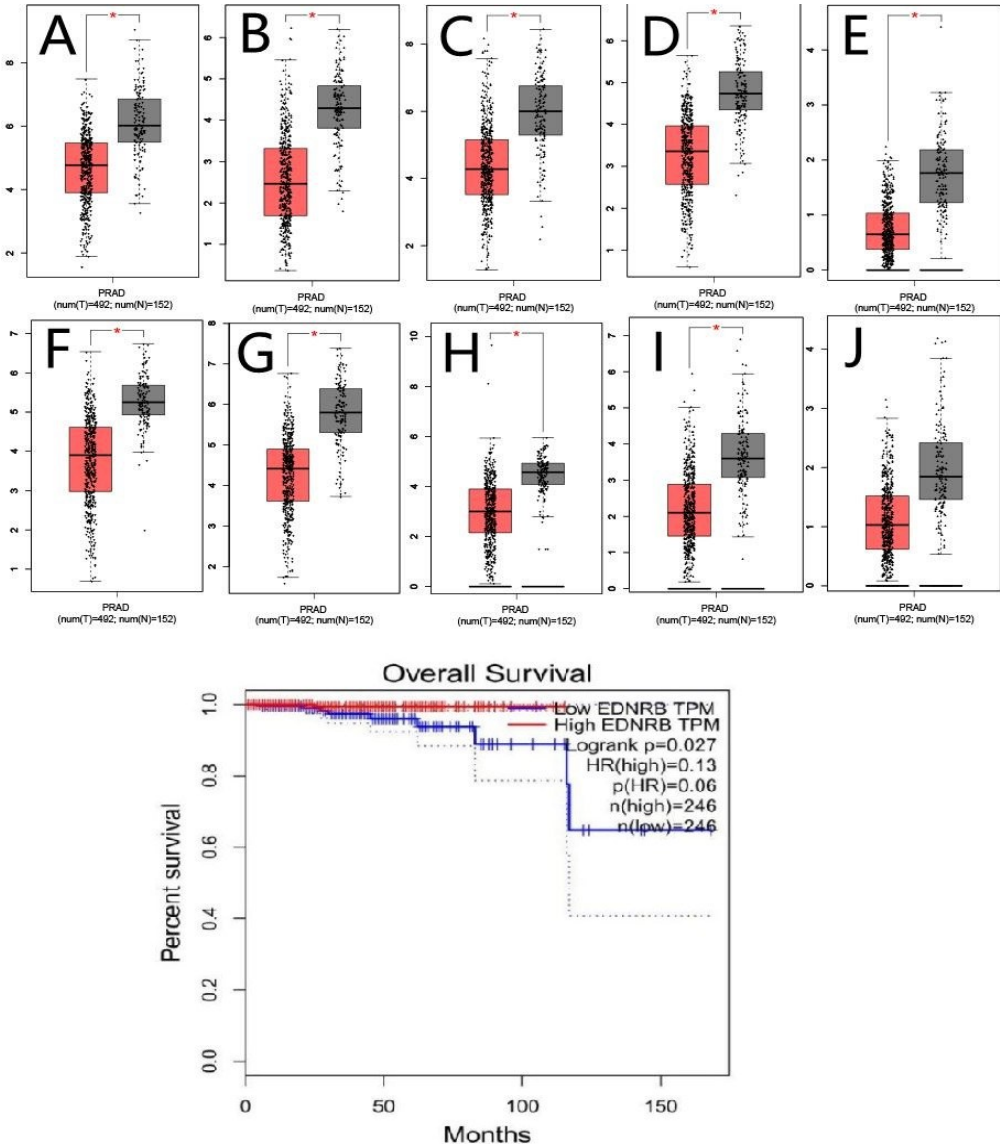


图 4 关键基因

2.4 表达与生存分析结果和验证 10 个关键基因在患者体内均为低表达,仅有 EDNRB 与患者的预后显著相关,且其低表达不利于 PCa 患者的总生存期(overall survival, OS) ($P < 0.05$),见图 5。根据 HPA 里下载的 TCGA 数据所作图验证 EDNRB 确在患者体内低表达且与预后有一定的相关性,见图 6。

3 讨论

有研究表明中国 2005—2014 年全国男性前列腺癌标化发病率上升 39.65%,城市发病率高于农村^[10],在 1992—2017 年间前列腺癌死亡率整体呈上升趋势,粗死亡率(1/100 000)从 1992 的 3.39 上升至 2017 的 7.17^[11]。前列腺癌与男性健康息息相关,而且由于其发病率和死亡率快速增加,预后较差,因此,应做好前列腺癌的筛查与预防。但早期前列腺癌同其他前列腺疾病诊断区分困难,前列腺癌防治仍是当今中国乃至全世界的公共卫生所要面临的重大难题。



注:A:CAV1,B:SNAI2,C:TGFB3,D:CAV2,E:PGR,F:MYOF,G:EHD2,H:FGFR2,I:EDNRB,J:PRKG1。

图 5 表达与生存分析

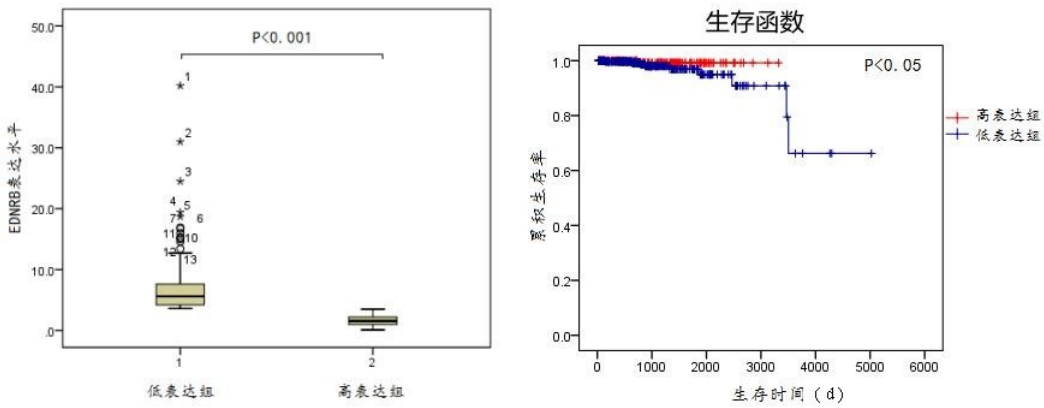


图 6 HPA 验证

中国抗癌协会发布的中国前列腺癌患者基因检测专家共识(2019 年版)^[12]中积极倡导组建泌尿肿瘤精准医学专家团队(genitourinary molecular tumor board, GU-MTB),设计精准化诊疗方案,建立由生物标志物引导的临床治疗路径,从而为其治疗提供更多选择。因此,从肿瘤大数据中筛查前列腺癌的关键基因,对其预防与治疗具有重要的意义。

本次研究共筛选出 103 个 DEGs,包括 34 个下调基因和 69 个上调基因。通过 GO 和 KEGG 通路富集分析发现这些基因富集在血管生成、脂质代谢等过程中,与 Hippo 信号等通路显著相关。早在 1976 年就有研究证明血管生成与肿瘤的进展关系紧密,尤其是与肿瘤细胞的增殖和转移有关^[13]。脂质代谢在肿瘤进展和转移侵袭等过程中也起到重要作用^[14]。目前已经有研究^[15-16]证实 Hippo 信号通路在肿瘤的发生发展中起关键作用,并且随着对 Hippo 信号通路的深入研究,发现其与心血管系统、免疫系统和其他系统也有重要关系。当 Hippo 信号通路发生了突变或者异常时,YAP/TAZ 开始过度表达且聚集在核中,其促进细胞增殖能力被激活,从而促进肿瘤发生发展。本研究发现 EDNRB(内皮素受体 B, endothelin receptor type B)基因的高表达患者生存率高于低表达患者,显示与前列腺癌的预后有意义。EDNRB 是一种内皮素受体相关抑癌基因,参与介导细胞信号转导,其启动子区发生甲基化而抑制其表达导致了肿瘤的发生^[17]。内皮素 B 受体可以通过促进细胞凋亡和清除 ET-1 来对抗肿瘤的进展^[18]。李志标等^[19]的研究也表明 EDNRB 与前列腺癌患者的生存率密切相关。同国内外研究对比,此次筛选出来的关键基因 EDNRB 与前列腺癌的发生发展过程确实有相关意义,有研究应用的意义。

综上所述,本研究利用多个数据库结合生物信息技术综合分析,发现 EDNRB 这个关键基因与前列

腺癌的预后相关,可能会为寻找前列腺癌潜在的生物标志物和研究其生物学机制提供一些有用的信息和方向,使后续进行的基础验证和临床研究更具有针对性,从而可以避免盲目的实验并且减少科研资源的浪费。但是生物信息学分析因为数据量、准确率等原因导致结果总会有一定的瑕疵,其结论最终均需要大量的实验和临床数据来验证。

参考文献:

- [1] Baciarello G, Gizzi M, Fizazi K. Advancing therapies in metastatic castration-resistant prostate cancer[J]. Expert Opin Pharmacother, 2018, 19(16): 1797-1804.
- [2] Siegel RL, Miller KD. Cancer statistics, 2020 [J]. CA Cancer J Clin, 2020, 70(1): 7-30.
- [3] Culp MB, Soerjomataram I, Efstathiou JA, et al. Recent global patterns in prostate cancer incidence and mortality rates[J]. Eur Urol, 2020, 77(1): 38-52.
- [4] Attard G, Parker C, Eeles RA, et al. Prostate cancer[J]. Lancet, 2016, 387(10013): 70-82.
- [5] Song EL, Song WT, Ren MH, et al. Identification of potential crucial genes associated with carcinogenesis of clear cell renal cell carcinoma[J]. J Cell Biochem, 2018, 119(7): 5163-5174.
- [6] Varambally S, Yu J, Laxman B, et al. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression[J]. Cancer Cell 2005, 8(5): 393-406.
- [7] Mortensen MM, Hoyer S, Lynnerup AS, et al. Expression profiling of prostate cancer tissue delineates genes associated with recurrence after prostatectomy[J]. Sci Rep, 2015, 5: 16018.
- [8] Arredouani MS, Lu B, Bhasin M, et al. Identification of the transcription factor single-minded homologue 2 as a potential biomarker and immunotherapy target in prostate cancer[J]. Clin Cancer Res, 2009, 15(18): 5794-5802.

- [9] Meller S, Meyer HA, Bethan B, et al. Integration of tissue metabolomics, transcriptomics and immunohistochemistry reveals ERG- and gleason score-specific metabolomic alterations in prostate cancer[J]. *Oncotarget*, 2016, 7(2): 1421-1438.
- [10] 刘灿, 李想, 王雷, 等. 2005—2014 年中国前列腺癌流行特征[J]. *中华疾病控制杂志*, 2021, 25(7): 806-811.
- [11] 丁雅妮, 刘晴, 张庆军. 1992—2017 年中国男性前列腺癌的死亡率趋势: APC 模型分析[J]. *公共卫生与预防医学*, 2020, 31(1): 25-28.
- [12] 叶定伟, 朱耀. 中国前列腺癌患者基因检测专家共识 (2019 年版)[J]. *中国癌症杂志*, 2019, 29(7): 553-560.
- [13] Folkman J, Cotran R. Relation of vascular proliferation to tumor growth[J]. *Int Rev Exp Pathol*, 1976, 16: 207-248.
- [14] Chen M, Zhang J, Sampieri K, et al. An aberrant SREBP-dependent lipogenic program promotes metastatic prostate cancer[J]. *Nat Genet*, 2018, 50(2): 206-218.
- [15] 刘金武, 张慧, 马宁宁, 等. Hippo 信号通路及相关疾病的研究进展[J]. *兰州大学学报(医学版)*, 2020, 46(1): 58-62.
- [16] Varelas X. The Hippo pathway effectors TAZ and YAP in development, homeostasis and disease[J]. *Development*, 2014, 141(8): 1614-1626.
- [17] Qureshi SA, Bashir MU, Yaqinudding A. Utility of DNA methylation markers for diagnosing cancer[J]. *Int J Surg*, 2010, 8(3): 194-198.
- [18] Bagnato A, Loizidou M, Pflug BR, et al. Role of the endothelin axis and its antagonists in the treatment of cancer[J]. *Br J Pharmacol*, 2011, 163(2): 220-233.
- [19] 李志标, 唐福才, 卢泽潮, 等. 基于大数据的前列腺癌生物信息学分析[J]. *中山大学学报(医学版)*, 2019, 40(6): 857-865.

收稿日期: 2021-09-01; 修回日期: 2021-09-12

(上接第 610 页)

- [5] Marquez R, Hantel A, Lorenz R, et al. A new family with a germline ANKRD26 mutation and predisposition to myeloid malignancies[J]. *Leuk Lymphoma*, 2014, 55(12): 2945-2946.
- [6] Pippucci T, Savoia A, Perrotta S, et al. Mutations in the 5' UTR of ANKRD26, the ankirin repeat domain 26 gene, cause an autosomal-dominant form of inherited thrombocytopenia, THC2[J]. *Am J Hum Genet*, 2011, 88(1): 115-120.
- [8] Ge ZH, Helmijr JCA, Jansen MPH, et al. Detection of oncogenic mutations in paired circulating tumor DNA and circulating tumor cells in patients with hepatocellular carcinoma[J]. *Transl Oncol*, 2021, 14(7): 101073.
- [9] 余结根, 李荣, 查晓娟, 等. 血清胃泌素及胃蛋白酶原在芜湖地区部分健康体检人群中的表达及意义[J]. *右江民族医学院学报*, 2020, 42(3): 308-311.
- [10] Bera TK, Liu XF, Yamada M, et al. A model for obesity and gigantism due to disruption of the ANKRD26 gene[J]. *Proc Natl Acad Sci USA*, 2008, 105(1): 270-275.
- [11] Hahn Y, Bera TK, Pastan IH, et al. Duplication and extensive remodeling shaped POTE family genes encoding proteins containing ankyrin repeat and coiled coil domains[J]. *Gene*, 2006, 366(2): 238-245.
- [12] Lei Y, Henderson BR, Emmanuel C, et al. Inhibition of ANKRD1 sensitizes human ovarian cancer cells to endoplasmic reticulum stress-induced apoptosis[J]. *Oncogene*, 2015, 34(4): 485-495.

收稿日期: 2021-04-26; 修回日期: 2021-05-23